

*Citation for published version:*  
Room, G 2014 'Evidence for Policy Making: Three Paradigms'.

*Publication date:*  
2014

[Link to publication](#)

**University of Bath**

### **Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# EVIDENCE FOR POLICY MAKING: THREE PARADIGMS

Graham Room, University of Bath

April 2014

## 1. INTRODUCTION

How can public policy-makers make good decisions? What counts as a good decision? And having made and implemented it, how can policy-makers check just how good it proved to be?

The most common answer nowadays is that policy decisions, to be good, should be evidence-based. First, it is only such policies that are likely to be effective. Second, with evidence to back them up, they can expect to command public support.<sup>1</sup>

Evidence-based policy-making (EBPM) became fashionable as part of the 'modernising' agenda with which Labour came into office in 1997. It was portrayed as the contemporary expression of the long-standing ambition, to bring scientific rationality to public affairs. For Labour, EBPM meant challenging established customs and vested interests with the demand for evidence of 'what works': it also meant abandoning ideology.

The Coalition government retained that commitment after 2010, with the launch of a network of ['What Works'](#) evidence centres for social policy. The Behavioural Insights Unit in the Cabinet Office has been playing a central role. One of their collaborators has been Ben Goldacre, well known through his *Guardian* columns on 'Bad Science'. He has in particular been a champion of randomised controlled trials (RCTs) in social policy. But the practitioner of good science does well to know and remember the history of his or her subject.

It seems often to be assumed that the application of RCTs to social policy is something new. Not so. At least as far back as Lyndon Johnson's 'Great Society' programmes of the 1960s, sociologists were grappling with the question: how could rigorous experimental design be applied to the testing of novel interventions - interventions which were taking place in the real world, in complex social and political contexts.

Marris and Rein (1974) provide the classic review of those efforts. Another of the US sociologists who got involved was Donald Campbell (1969). In subsequent years he provided one of the standard texts on 'quasi-experimentation': how to remain rigorous even under the limitations of practical research and evaluation (Cook and Campbell, 1979). Years later it was after him that the [Campbell Collaboration](#) was named: applying to social policy interventions the protocols of systematic review that the [Cochrane Collaboration](#) had pioneered in the medical field.

Good science also means thinking about the variety of concepts and methods to hand, and which to use in different situations. Which concepts and methods work best in relation to which research questions and problems? This is a form of meta-thinking. What follows suggests three alternative paradigms for assessing 'what works'.

## 2. EBPM AS THE DISAGGREGATION OF IMPACT

The advocates of EBPM have been concerned, first and foremost, with evidence of the outcome or *impact* of a particular intervention. Evidence is collected, evaluated and aggregated –

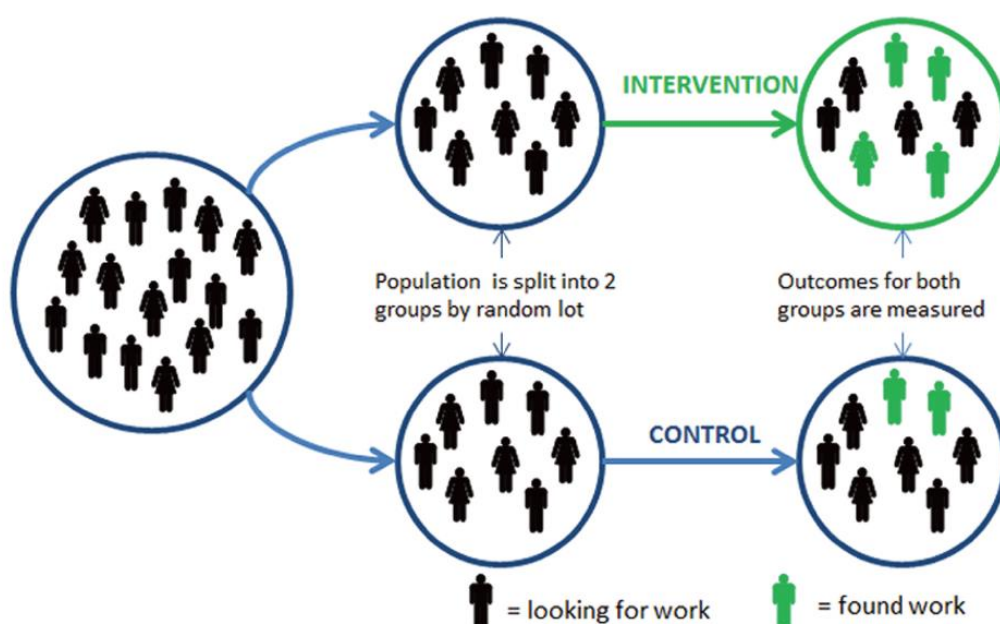
---

<sup>1</sup> The alternative to evidence-based policy making – and the implied target of its critical thrust – is not always clear. It certainly however encompasses policies and practices that reflect political loyalties, professional self-aggrandisement or the pressure of lobbyists.

'systematically reviewed' - across as wide a range of contexts as possible. This is meant to produce a rigorous assessment of 'what works' - and in some degree an understanding of how delivery can be adjusted to a variety of conditions.

The gold standard of EBPM remains the randomised controlled trial. It is to this that policy makers and their public critics regularly appeal (see for example HM Treasury (2011) para 9.16; Johnson (2011)). There is a well-defined intervention which has a measurable impact on the target population in question. This intervention is administered to a 'treatment' group and the effects are compared with those for a 'control' group. A clear and straightforward example is provided in the Cabinet Office document, *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials* (Haynes et al., 2012), from which **Figure 1** is taken.

**Figure 1: The Basic Design of a Randomised Controlled Trial (RCT)**

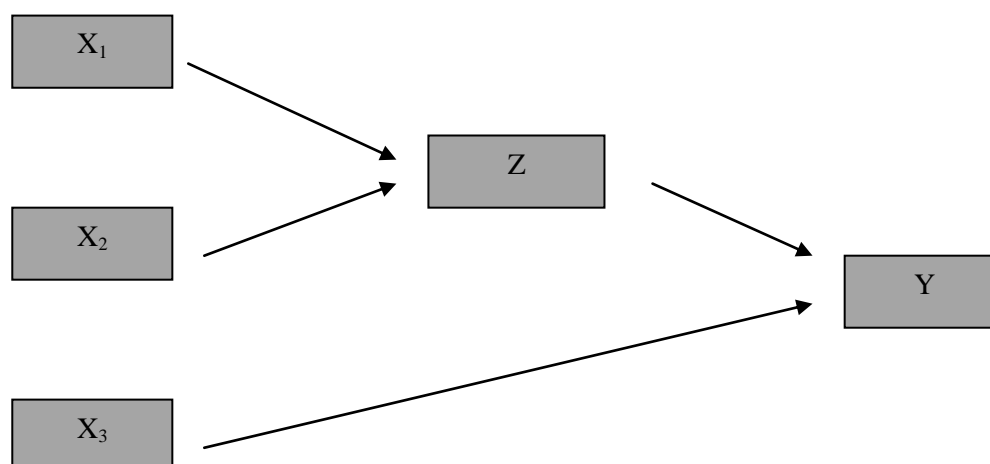


This approach to EBPM is variously elaborated and criticised in the policy literature. Much of the discussion involves relaxing the severe demands of the RCT, successively abandoning one or more of its defining elements and making do with quasi-experimental and descriptive studies (Pawson, 2006: Box 3.2). Campbell himself was central to the development of such 'softer' methodologies and the assessment of their utility. Even then however, some minimum threshold of rigour must be retained, so as to be able to establish 'what works' with sufficient certainty.

In practice of course any new policy is launched into a world already crowded with policy initiatives, ancient and modern, whose effects and impacts interact. Nevertheless, if we make some simplifying assumptions, appropriate statistical methods are available with which we can, in principle, partition and disaggregate these effects and isolate the contribution of any particular intervention. Such methods can thus disentangle the combined effects of multiple interventions, against the background of a changing environment. That, at least, is what is commonly maintained (see for example Harkness et al (2009)).

We might capture these key elements of EBPM through **Figure 2**. The independent variables  $X_1$ ,  $X_2$  and  $X_3$  correspond to the interventions that are simultaneously under way and that (in part through their effect on  $Z$ ) affect the impact variable  $Y$ . Each of these can vary (albeit within some bounded range), in ways that our systematic review of the evidence will reveal.

**Figure 2: Policy Intervention as Additive Impact**



The real world is rarely so simple. Variables exert their effects within different timescales; there may be threshold and ratchet effects; impact may not increase in strict proportion to the independent variables. Econometric techniques exist for handling some of these complications, so that it is still possible to separate out the effects of these various interventions (Room and Brown, 2013). Nevertheless, to separate them in this way is more than just a technical matter. It carries an implicit ontology of the social world, as one that can be disaggregated into a set of independent ‘variables’ that additively compose this world’s causal mechanisms. Epistemologically sophisticated methodologies are here being used to trump ontological complexity.

### 3. THE ONTOLOGICAL CHALLENGE OF REALISM

Pawson (2006) questions the language of ‘impact’ as far as social policy interventions are concerned. He insists that such interventions do not so much ‘impact’ upon social actors as ‘engage’ with them: both the ‘street level bureaucrats’ who deliver the interventions and the members of the target population. Such actors learn by doing; if allowed to do so, and given a degree of freedom, they may well improve on the policy maker’s design. But of course, they also have their own agendas: they may contest the goals of the intervention and subvert it to their own ends.

The intervention in question is therefore likely to take different forms in the hands of different stakeholders and in different institutional contexts: we can hardly speak of the ‘standard’ form of the intervention and its effectiveness, without regard to these active subjects. Evaluation of the intervention therefore involves laying out the variety of forms that it takes – and with what effects – under these different contingencies.

On the basis of this critique, Pawson offers a response in terms of ‘realism’, as developed by philosophers of science such as Harré (1972: Ch 4). Realism insists that it is not enough to establish by appropriate statistical techniques the correlations of independent and dependent variables – the central concern of Figure 2. Explanation must also include an account of the real world processes which produce and underpin the patterns we observe.<sup>2</sup> Ontology matters: it cannot be entirely dodged by using epistemologically sophisticated methodologies.

---

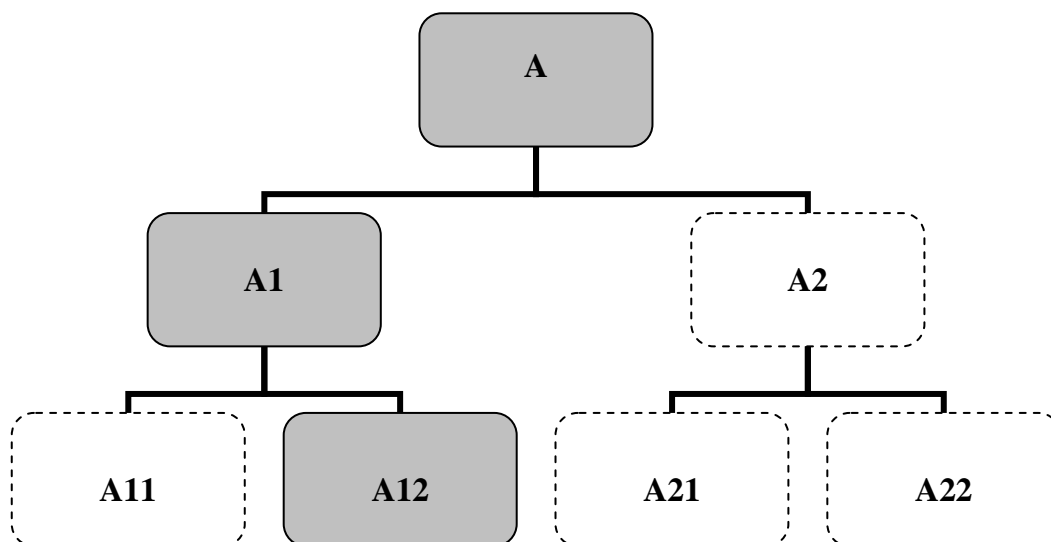
<sup>2</sup> Pawson repeatedly speaks of this as the development of theory and models – our imagining of the generative processes which underlie the patterns we observe. These theories then guide further investigation, in the course of which they are

For Harré, these ‘generative mechanisms’ involve potentialities that are unlocked or closed down by different contextual conditions. Explosives such as gunpowder and dynamite provide the example that Harré and his followers commonly cite, including Pawson. The chemical composition of the explosive provides the capacity to explode: but whether it does so or not depends on such factors as the absence of damp, the presence of oxygen, the ambient temperature, etc. Causal analysis of generative mechanisms and policy impacts must be alert to such contingencies.

Pawson spells out the methodological implications. Causal analysis is an ‘explanatory quest’, whose goal is to disentangle and peel away these contingencies. As we do this, we will develop a cumulative understanding of the interventions in question and how they are likely to behave in different situations.

**Figure 3** offers a visualisation of this process (albeit the diagram is ours, not Pawson’s). An intervention **A** can, depending on the conditions of its implementation, take form **A1** or **A2**; and depending on further contingencies, may then take the form **A11**, **A12**, **A21** or **A22**. The Figure tells us that **A12** was the form taken in a given instance; but asks us, in a realist spirit, to peel away and reveal the contingencies that came successively into play and the other possibilities that might have been realised.

**Figure 3: Policy Intervention as Contingent Diversification**



Nevertheless, this still leaves Pawson focussing primarily on the individual intervention – whether gunpowder or a new pharmaceutical product or a social policy programme – and unpicking the contingent factors that activate or inhibit its impact. We now seek to go beyond this, to the crowded real world of multiple and interacting policy initiatives.

#### 4. THE CASE FOR TRANSFORMATIVE REALISM

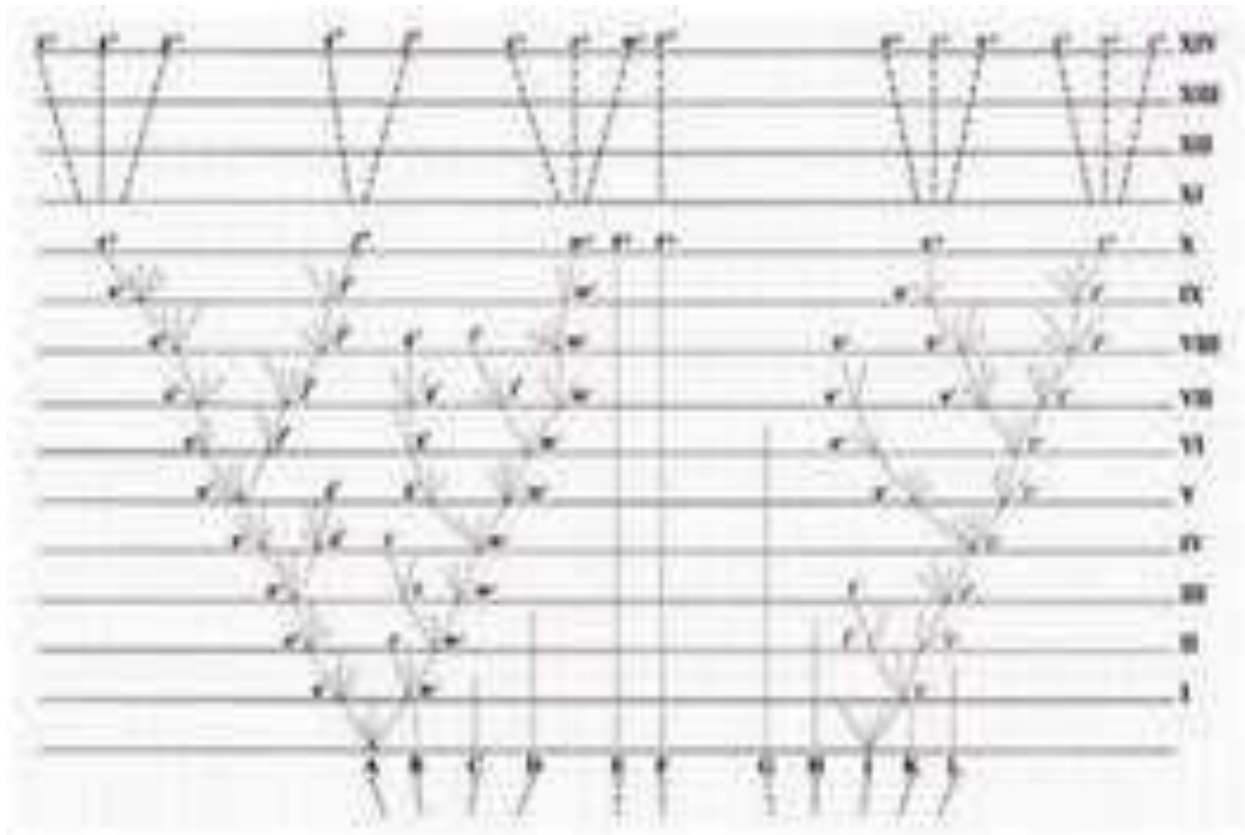
Figure 3 is a tree diagram – a dendrogram – with successively sprouting branches and sub-branches. It is therefore reminiscent of Darwin’s ‘Tree of Life’, as reproduced in **Figure 4** (Darwin, 1859). In his account of the diversification of species, Darwin was centrally concerned with

---

liable themselves to be modified (page 100). They also provide a mental model of the world, by which decision-makers can navigate the ‘tortuous pathways’ along which they travel (page 170).

processes of adaptation to the successive contingencies of different habitats, just as Pawson's realism exposes the successive contingencies which shape a policy intervention and its effects.

**Figure 4: Darwin's Tree of Life**



Nevertheless, Darwin also referred to the co-evolution of species: albeit not perhaps to the same extent as his successors, who have shown how powerfully the dynamic synergies of co-evolution shape the evolutionary story (Kauffman, 1993; Maynard Smith and Szathmary, 2000). Such co-evolution typically involves populations that are far removed from each other in the evolutionary tree: for example, flowers and insects, mutually favouring each other's 'struggle for existence' over the last 140 million years.

It is interactions of this sort, among policy interventions, that we now bring centre-stage. Both of the previous paradigms assumed that we can focus on a single policy intervention in isolation. But in the real world, any intervention unfolds not on a *tabula rasa*, but within a dynamic policy 'ecosystem'. The policy maker needs to be able to anticipate such dynamic effects – and to judge which ones will accelerate and reinforce his or her policy ambition, and which ones throw it off course. This is critical for any assessment of 'what works' (Room, 2013).

Policy interventions are launched into a crowded world. These forerunners are not the mere detritus of policy enthusiasms long forgotten; in many cases their champions are still at work, seeking to broaden their scope and colonise the landscape onto which any new policy is launched. Previous interventions shape the fears and hopes and expectations with which the public view the new intervention. Around them constituencies and vested interests will have formed that may favour or oppose the new intervention. More than this, the new intervention is liable to trigger dynamic synergies with some elements of the policy system – forms of 'co-evolution' which accelerate changes in direction which cannot be understood as the simple consequence or impact of the new intervention. Equally however, the new intervention may be unable to break into policy ecosystems that are resilient against such new 'invaders'.

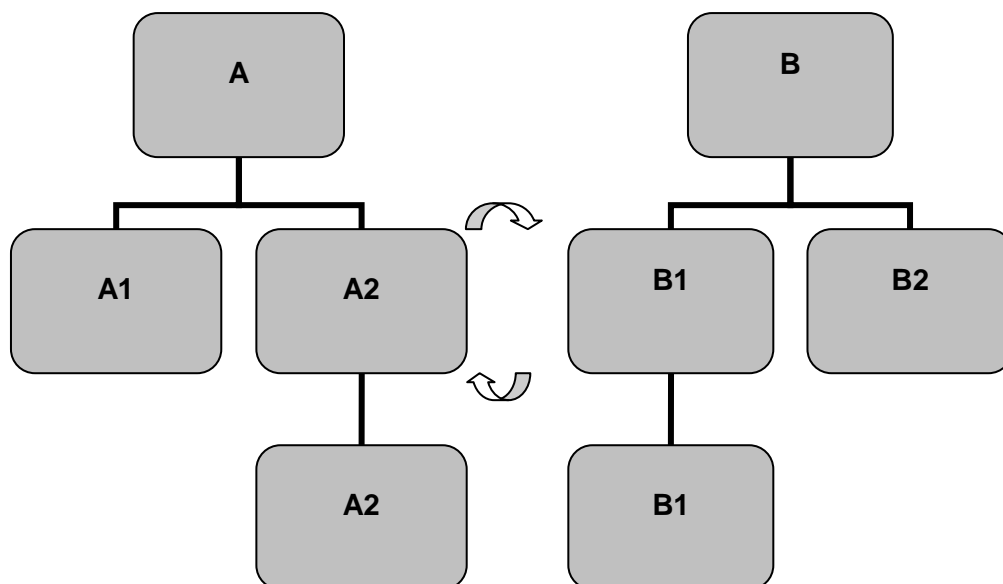
It is not just a matter of what policies co-habit a given landscape. What can also matter is the order in which they have been introduced. Sequence and timing are important: change them, and the ways in which they shape each other will also change. Not only must the search for evidence regarding a policy intervention consider how it will work in combination with other policies; it must also differentiate according to the order in which those policies are introduced, and having regard to the different time scales of their likely effects.<sup>3</sup>

This also means that policy interventions and their potentialities are not fixed, in the sense that the chemical composition of gunpowder is fixed. We are interested not in gunpowder *per se*, but in the weapons technologies of which it is no more than a component, and whose potentialities, far from fixed, will then be the stuff of desperate arms races. This is an *evolutionary* version of realism. The focus is still on 'generative mechanisms', but these are now located not so much within individual interventions, but rather in the transformative synergies that develop *among* these interventions and their stakeholders and by which they co-evolve.

To use the term 'evolutionary' perhaps holds risks of misunderstanding. Evolution by natural selection is a blind process. In human societies in contrast, people in some degree make their own history. They probe and they experiment, not randomly but by systematic testing and learning (Bronowski, 1981: Chs 2-4). They re-shape the technologies and institutions of their world, in hope of discovering new dynamic synergies from which they can benefit. They strive to develop thereby their understanding and their capacities; their control over their lives; their positional advantage and leverage. This brings interests and power and politics centre-stage. We will therefore henceforth speak not of evolutionary but of *transformative realism*.

As we have seen, visual representations can provide powerful images that organise and direct our thinking. Figures 2 and 3 provided such images for our first two paradigms. We now seek a counterpart for transformative realism, in its simplest or canonical form.

**Figure 5: Policy Intervention as Transformative Synergy**



<sup>3</sup> Thus for example western advice to Russia in the 1990s, to privatise enterprises even before well-functioning markets and financial institutions had been established, is now widely reckoned to have had damaging consequences for Russian economy and society. Or to take a very different example, popular perceptions of health hazards in light of previous health panics - and the way that policy makers have handled them - have consequences for subsequent policy interventions, and for the sequencing of communications about such risks (Breakwell, 2001; Barnett and Breakwell, 2003).

In **Figure 5**, **A** and **B** are two policy interventions among many. Each may have been well-specified by the instigators; nevertheless, each involves multiple layers of policy staff, bringing a diversity of interpretations and interests to bear, as Pawson argues. This process of diversification we represent by the variations **A1** and **A2**, **B1** and **B2**.

No intervention however is isolated: each interacts with others. What now matters is which of the four sets of interactions between **A1** and **A2** on the one hand, **B1** and **B2** on the other, produces the most powerful transformative synergies. In the diagram, we show the relationship of **A2** with **B1** as being this favoured pairing, this 'elective affinity'. **A2** and **B1** will each now accelerate the flourishing of the other: they progressively dominate **A1** and **B2**, which are re-ordered, marginalised, frozen or extinguished altogether.

Thus by the time we arrive at the time period represented by the bottom row of the diagram, **A2** and **B1** dominate. This is a policy world substantially different from the one with which we began, centred on **A** and **B**. Nevertheless, domination by **A2** and **B1** will not last for ever; further rounds of interaction with the larger policy 'eco-system' will eventually destabilise them, as new rounds of variation and selection are set in motion. In these new rounds, **A1** and **B2** will however no longer be in play, or they will at least have been marginalised: there will be little chance for their potential synergies with new partners to be tested. Sequence matters, because it dictates which elements are made available for subsequent interaction with others.

This is how we may visually represent the third of our paradigms of evidence-based policy making, in terms of 'transformative realism'. The task of the policy analyst is to identify the dynamic synergies by which **A** and **B** and their sub-variants interact with each other and thereby come to dominate the changing morphology of the system as a whole. Or, if we take **A** as the policy environment that exists initially, we seek to understand the consequences of a new policy intervention **B** 'invading' the system and re-sculpting that environment; or alternatively, revealing that this environment is sufficiently stable and resilient against such invasion as to remain in its original state.

## 5. CHOOSING A PARADIGM

The ontological objections we have raised do not mean the wholesale rejection of EBPM as the assessment of impact. Nor do they necessarily mean the entire abandonment of the randomised controlled trial. It is a matter of practical judgment, how far these procedures can still provide useful guidance, in particular empirical situations.

Figure 2 will often therefore remain a valuable point of reference. The multiple contingencies of Figure 3 and the transformative synergies of Figure 5 are not all-pervasive. Some degree of uniformity and stability are preconditions of all policy-making. What must not however be overlooked is that they are contingent.

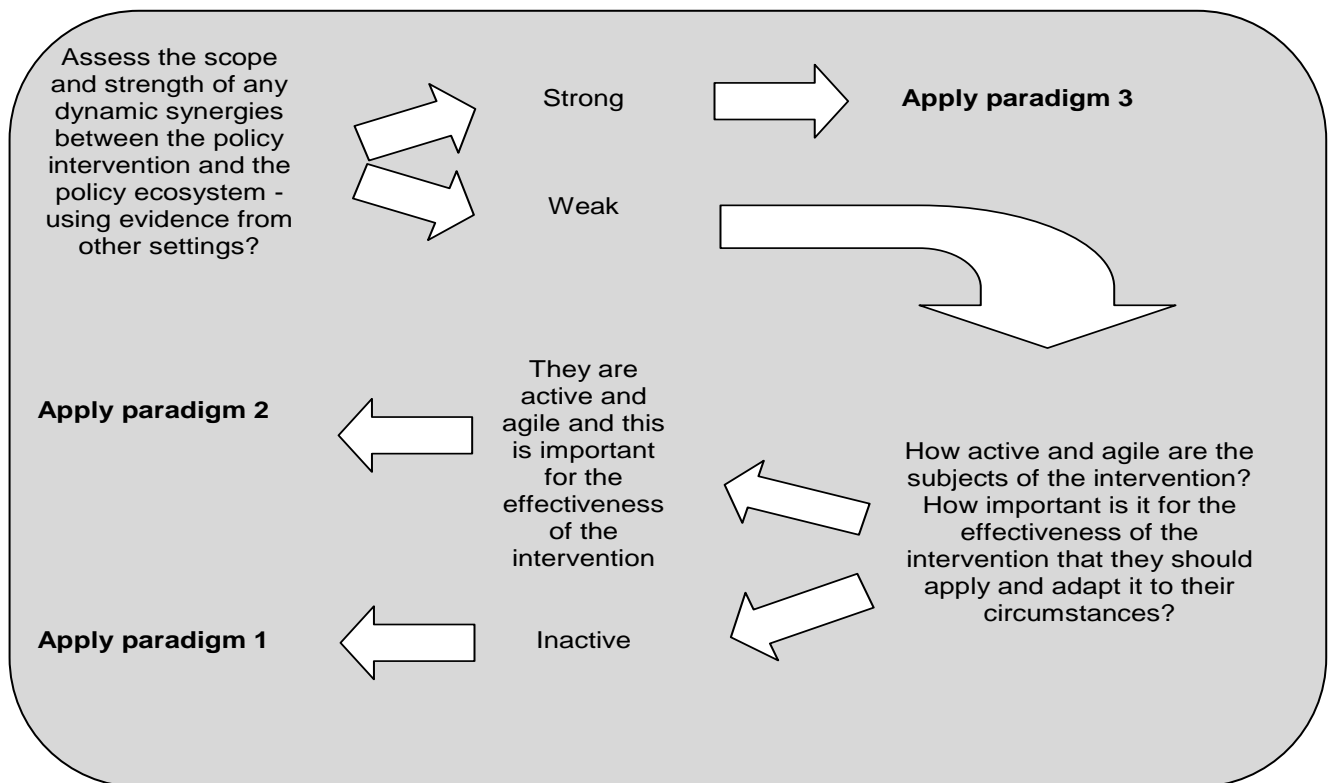
How should the policy analyst choose which paradigm is appropriate? **Figure 6** provides a suggested procedure for making that choice.

This choice is not however just a technical question. It involves judgements as to the significance of different dynamic synergies, in relation to the objectives not only of policy makers but also of other stakeholders across the communities affected.

Policy making is a contested process, unavoidably involving interests and power and politics. Intrinsic to such struggles is the very definition of different societal 'problems'. Who is to be blamed for these problems and how far is there a responsibility on the public authorities to address them (Butler and Drakeford, 2005)? What standards of evidence are demanded for different problems, as a precondition for the investment of public resources? Which problems require a novel response - and when is such novelty no more than a way to avoid hard political questions?



**Figure 6: Choosing a Paradigm**



The struggle is moreover not just for resources and position but also over the very way that we 'see' the world, both as it is and as it ought to be. It is therefore, not least, a cultural struggle, over the legitimating symbols that give stability to our social world and the cultural hegemony of powerful groups.

It is on just such a stage that policy analysts attempt to develop an evidence base for policy and practice. What they provide must therefore take full account of the political economy and distribution of power within which struggles over the future of the social and political order are being waged. If they sanitise and cloud this task, in the language of technical measurement and reified system dynamics, this is itself a political choice.

**Acknowledgements:** I am grateful to a number of colleagues for comments during the preparation of this paper. It draws upon a programme of work supported by the Economic and Social Research Council (Award RES-063-27-0130) and my book *Complexity, Institutions and Public Policy: Agile Decision-Making in a Turbulent World*, Cheltenham: Edward Elgar (2011). It draws upon an earlier and more extensive paper, published as Room, G. (2013), 'Evidence for Agile Policy Makers: The Contribution of Transformative Realism', *Evidence and Policy*, 9(2): 225-244.

## REFERENCES

Barnett, J. and G. Breakwell (2003), 'The Social Amplification of Risk and the Hazard Sequence: the October 1995 Oral Contraceptive Pill Scare', *Health, Risk and Society*, 5(3): 301-313.

- Breakwell, G. (2001), 'Mental Models and Social Representations of Hazards: the Significance of Identity Processes ', *Journal of Risk Research*, 4(5): 341-351.
- Bronowski, J. (1981), *The Ascent of Man*, London: Futura
- Campbell, D. T. (1969), 'Reforms as Experiments', *American Psychologist*, 24(4): 409-429.
- Cook, T. D. and D. T. Campbell (1979), *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Chicago: Rand-McNally
- Darwin, C. (1859), *The Origin of Species*, London: Wordsworth (reprinted 1998)
- H M Treasury (2011), *The Magenta Book: Guidance for Evaluation*, London: HM Treasury
- Harkness, S., P. Gregg, and S. Smith (2009), 'Welfare reform and Lone Parents in the UK', *The Economic Journal*, 119(535): 38-65.
- Harré, R. (1972), *The Philosophies of Science*, Oxford: Oxford University Press
- Haynes, L., O. Service, B. Goldacre, and D. Torgerson (2012), *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials* London: Cabinet Office Behavioural Insights Team.
- Johnson, P. (2011), 'New policies, like new medicines, should first be put to the test'. *Guardian*. Tuesday 27 September.
- Marris, P. and M. Rein (1974), *Dilemmas of Social Reform*, Harmondsworth: Penguin
- Pawson, R. (2006), *Evidence-Based Policy: A Realist Perspective*, London: Sage
- Room, G. (2011), *Complexity, Institutions and Public Policy: Agile Decision-Making in a Turbulent World*, Cheltenham: Edward Elgar
- Room, G. (2013), 'Evidence for Agile Policy Makers: The Contribution of Transformative Realism', *Evidence and Policy*, 9(2): 225-244.
- Room, G. and G. Brown (2013), 'Contingent Development and Causal Explanation', (*forthcoming*).